



# Proximal Deixis with Calendar Terms: Cross-linguistic Patterns of Temporal Reference

Dana Cohen, Sylviane R. Schwer

## ► To cite this version:

Dana Cohen, Sylviane R. Schwer. Proximal Deixis with Calendar Terms: Cross-linguistic Patterns of Temporal Reference. 2011. halshs-01059978

**HAL Id: halshs-01059978**

**<https://shs.hal.science/halshs-01059978>**

Preprint submitted on 23 Jun 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Proximal Deixis with Calendar Terms: Cross-linguistic Patterns of Temporal Reference<sup>\*</sup>

**Dana Cohen** (SFL, U. of Paris 8)  
dcohleng@inbox.com

**Sylviane R. Schwer** (LIPN, U. of Paris 13)  
Sylviane.Schwer@lipn.univ-paris13.fr

This paper provides a comparative perspective of proximal deictic temporal markers, examining both language-specific and cross-linguistic patterns. The expressions considered are based on three types of time-reckoning terms, reflecting distinct types of abstract segmentation and referring to distinct temporal cycles: FUNDAMENTAL CALENDAR UNITS (e.g., day, year), DAY SEGMENTS (e.g., morning, noon) and NAMED DAYS OF THE WEEK (e.g., Tuesday). The results indicate that such temporal expressions are not homogeneous, pointing to a number of factors that affect their properties and the retrieval of temporal reference: the type of temporal cycle and segmentation used, the transparency or opacity of the information signalled, distance from the deictic centre, (in)dependence of the expressions as temporal markers, and relations between the relevant temporal segmentations.

**Keywords:** temporal reference, proximal deixis, temporal segmentation, time reckoning, calendar terms

## 1. Introduction<sup>1</sup>

This paper reports initial results of an ongoing cross-linguistic study of temporal markers that include time-reckoning terms, particularly those designating conventionalised segmentations of clock-and-calendar systems (such as *day*, *month*, *weekend*, *winter*). The linguistic encoding of temporality involves a range of features and structures that reflects a combination of linguistic and non-linguistic factors, perceptual, social and cultural. This is particularly true when dealing with conventionalised time-reckoning terms. Our study therefore aims to distinguish the linguistic and non-linguistic factors involved and determine their impact on the structure and use of the relevant temporal markers.

---

<sup>\*</sup> We would like to extend our warm thanks to all our colleagues, friends and informants, who supplied us with data and judgments on their own languages. An earlier version was presented at the TEMPTYPAC Workshop (Paris 11-12 March 2010). We would like to thank the participants for their instructive comments and discussion. The initial part of this study was made possible thanks to a short-term research fellowship granted to D. Cohen at LIPN in 2009.

<sup>1</sup> Abbreviations in the glosses: DF = definite; EM = existential marker; F = fem.; M = masc.; N = neuter; PL = plural; POSS = possessive.

In this paper, we focus on the proximal deictic use of three types of expressions employing time-reckoning terms: FUNDAMENTAL CALENDAR UNITS (e.g., day, year, week, month), DAY SEGMENTS (e.g., morning, noon, night) and NAMED DAYS OF THE WEEK (e.g., Tuesday). These reflect distinct types of temporal segmentation and refer to distinct temporal cycles, factors which will be shown to bear on their linguistic behaviour. Our study takes a closer look at the form of such temporal markers from a cross-linguistic perspective, examining their internal structure and the temporal information they encode, taking into account the formal properties of the abstract segmentations they employ. To our knowledge, no other study takes the same approach.

Temporality has been extensively studied in linguistics as it has been in many other domains, from physics to philosophy. In the linguistic literature, the vast majority of studies are centred on phenomena related to the predicate, the interactions of tense, aspect, and modality (some notable studies in English: Jespersen, 1924; Reichenbach, 1947; Bach, 1981; Kamp & Rohrer, 1983; Comrie, 1985; Smith, 1991, 2007; Hornstein, 1993; ter Meulen, 1995; Abusch, 1997; Zagana, 2007; and in French: Beauzée, 1767; Guillaume, 1929; Wilmet, 1976; Vet, 1980; Gosselin, 1986; Saussure, 2003). Such studies generally focus on the sequencing of predicates and the placement of different events along a timeline. In such research contexts, expressions like *tomorrow*, *last year*, *for two hours* are often taken as direct temporal signals, providing explicit anchoring but not analysed themselves. Other studies focus on specific temporal expressions, such as adverbs (e.g., *already*, *still*), prepositions (*avant/après*), or nominals (*l'an dernier*), primarily from a language-specific perspective (Borillo, 1983, 2005; Berthonneau, 1989, 1990; Kleiber, 1993; Haspelmath, 1997; Giorgi & Pianesi, 2003; José, 2003; Bras, 2005; Demirdache & Uribe-Etxebarria, 2007; Giorgi, 2008; *inter alia*). A number of analyses consider the interaction of tense, aspect and mood with temporal adverbials cross-linguistically (Vendler, 1967; Dowty, 1982; Chung & Timberlake, 1985; Giorgi & Pianesi, 1997; Bhat, 1999; *inter alia*), but these again focus only on the predicative perspective. In particular, the study of the deictic properties of tense (past / present / future, and proximity)<sup>2</sup> has not been accompanied by a systematic examination of the deictic series associated with time-reckoning terms.

A parallel problem is evident in the domain of natural language processing, which has seen many attempts to employ formalised calendar systems (such as Cukierman & Delgrande, 1998; Schwer, 2002). The aim of most annotation systems of temporal expressions is, again, the anchoring of events along a timeline and the

---

<sup>2</sup> Deictic marking of distance from a centre is well-established, with deictic determiners or in the spatial domain (e.g., Filmore, 1997). Non-numerical spatial distance is often lexicalised cross-linguistically, frequently in a binary proximal/non-proximal division, (Eng. *here/ there*, Fr. *ici/ là/ là-bas*, Gr. *hier/ da/ dort*), while the medial vs. distal distinction is less often lexicalised as a separate root (as in Latin proximal *h-*, medial *ist-*, distal *ill-*, and indefinite *ali-* or Vietnamese *dây/ do/ kia* (Nguyen, 1992)).

Tense systems have been studied in terms of non-numerical distance along similar lines (Dahl, 1983, 1985; Comrie, 1985). The past tense distinction in 17th century French (Arnauld and Lancelot, 1660) has been analysed in terms of the well-known *règle des 24 heures*, drawing a binary distinction, hodiernal vs. non-hodiernal (passé composé *j'ai écrit cette lettre ce matin* 'I wrote this letter this morning' vs. passé simple *j'écrivis cette lettre hier* 'I wrote this letter yesterday'). Other languages, (e.g., Chibemba (Bantu: Givon, 1972) and Maba languages (Bonomo, 2001)) offer finer distance distinctions, with at least a three-level partition in each direction: hodiernal/ hesternal/ prehesternal.

evaluation of relative durations by linking time-reckoning expressions into a calendar or some chronoline, here via the automatic identification of such expressions. In order to identify temporal expressions based on calendar units, some dating grammars have been developed in some languages (e.g., Gross, 2002 for French, Flickinger, 1996 for English). The anchoring task is generally achieved through transforming time-reckoning expressions into a standard format – usually ISO 8601 standard – (Pustejovsky et al., 2005a, 2005b). This identification task, although taken to be trivial, encounters difficulties, raising the need for a different approach and a detailed study of non-verbal temporal expressions (Battistelli et al. 2006; Mazur & Dale, 2008; Cailliau et al., 2009).

Another aspect that is relevant for the present study has to do with the complexity accorded to the abstract representation of time and to the relations between the various segmentations. Temporality studies generally conceptualise the timeline along which events are sequenced as one-dimensional, a linear horizontal axis (see, Girard 1747; Beauzée, 1767; Reichenbach, 1947; Comrie, 1985; Smith, 1991; Haspelmath, 1997; *inter alia.*). Yet, several researchers have argued that linguistic time is multi-dimensional (see, Guillaume, 1929; Bull, 1960, Benveniste, 1974; Traugott, 1978; Evans, 2005; Botne & Kershner, 2008). Mathematical models of calendars tend to base the segmentation of the timeline on a base segment (usually the second), which serves to define all other segmentations, assuming a hierarchical organisation of temporal units. However, this approach does not address differences between types of segmentations in calendar systems and relations which go beyond hierarchical inclusion and granularity. Thus, they do not take into consideration the potential impact that different segmentations and their properties may have on the linguistic structures.

Temporal indicators that lie outside the verbal category, most of which involve time-reckoning terms, have been the subject of few studies. As noted above, such markers tend to be seen as signalling their reference directly and explicitly. The difficulties encountered by automated models indicate that these apparently simple, explicit markers require further study. Our focus on these non-verbal temporal markers thus serves as a complementary perspective to the predicate-oriented approach. Our primary aim is to identify and characterise cross-linguistic patterns along three aspects: the structural form of the temporal markers, their referential properties, and the properties of the abstract segmentations involved.

As will be shown, the expressions examined reveal systematic linguistic patterns in the temporal information they signal, in the way this information is encoded, and in their ability to function independently as temporal markers. The conventionalisation of temporal expressions using time-reckoning terms involves several inter-related aspects: the transparency or opacity of temporal information signalled by the expression, distance from the deictic centre, (in)dependence of the expressions as markers, the type of abstract time-reckoning segmentation involved and the relations between the relevant temporal segmentations.

Temporal expressions involving time-reckoning terms are traditionally categorised as adverbs in various languages. However, while their function in the clause may be adverbial, the distribution and structure of these expressions in the set of languages examined clearly indicate a nominal categorisation and internal structure, exhibiting plurality, gender, case, and appearing with demonstratives,

definite determiners, possessive markers, quantification, and prepositions.<sup>3</sup> Consequently, we consider these temporal markers to be nominal expressions; the calendar term heads a nominal phrase, whatever the function of this phrase is in the clause.<sup>4</sup> On par with other nominal forms, we assume that the temporal markers under discussion are referential elements, signalling to the addressee a temporal fix to be identified, which is frequently anchored to another temporal reference point. Our study indicates that the necessary information for reference resolution is encoded in different ways. Of particular interest is the fact that temporal reference cannot always be achieved through such markers independently of other elements. We therefore wish to establish the extent of this phenomenon—how much of the necessary information is directly encoded in these expressions and how much is supplied by additional linguistic and extra-linguistic information.

This article presents initial results of an ongoing cross-linguistic study. The fundamental units have been examined (to varying degrees of depth) in over 20 languages of distinct families. Other segmentations have been examined in a more limited set of languages, so far. The data for this study combine free production and elicited input, as well as attested data from written sources for some of the languages.

Formal temporal reference requires three components of information: the size of the temporal segment (given the relevant calendar), the distance from the central anchor, and the orientation (future/Ø/past). The deictic expressions examined differ in whether all three components are present, and the degree of their transparency, that is, the degree to which they are discernable as distinct elements, ranging from the clearly compositional (*2 days ago*) to the completely opaque (Fr. *hier* 'yesterday'). Distance from the deictic centre affects the degree of opacity and the degree of specialisation of the linguistic sequence: distal reference employs free and productive strings, typically sequential and transparently compositional, whereas reference to the proximal and central zones employs specialised constructions that are, frequently, increasingly opaque and conventionalised in form and reference. These temporal expressions exhibit distinctions in their ability to signal temporal reference independently. Some expressions cannot achieve accurate temporal reference without additional temporal cues from the rest of the clause or extra-linguistic information. Thus, as shown below, the temporal reference of named days of the week (*Tuesday, next/last Tuesday*) relies heavily on the identity of the day of speech (Day S) and its distance to or from the named day in the linguistic string. Accurate reference using the day segments can only be established through information provided by the rest of the clause, notably the tense. Finally, while the relations between calendar units are based on the relevant abstract timekeeping system (in terms of granularity, inclusion and sequence), the significance of these relations in

---

<sup>3</sup> Even the least modified of these expressions productively exhibit such nominal features (Eng. *yesterdays, every yesterday/tomorrow*; Fr. *tous les aujourd'hui* (every DF.PL today.PL); It. *ogni ieri* (every yesterday); Heb. *etmoleinu* (yesterday.PL.POSS)).

<sup>4</sup> Wilmet (2003: 467) classifies the French expressions *aujourd'hui, hier* and *demain* (today, yesterday, tomorrow) as temporal indefinite pronouns.

linguistic calendar expressions varies, minimally affecting their degree of independence and the additional information required (thus, in French, use of named days is affected by the identity of Day S, but not by the containing segment Week S).

As a starting point, we highlight key factors of the abstract construct of calendars from a configurational computative perspective and relating them to the linguistic tools (section 2). The linguistic use of the three types of expressions is outlined in sections 3–5, highlighting the emerging cross-linguistic patterns. Section 6 considers the consequences of these patterns from a mathematical and computational perspective.

## 2. Timekeeping Systems

The notion of time encompasses a cognitive perception affected by cultural factors (and, in some societies, correlated with a formal mathematical concept). Thus, the human perception of time includes two fundamental opposing relations, succession and simultaneity (themselves linked to mathematical enumeration, ordering, and magnitudes), as well as to the notions of past, present, and future, proximity and distance, which are related to deixis. The abstract time-keeping system that is the calendar provides a set of divisions, organisation, and naming of temporal periods that is intended to supply a practical measurement of time to coordinate human activity and maintain social function. The principal of a calendar is the division of the abstract, metaphorical image of time into more or less identical segments (fig. 1), so that every punctual event can be assigned to a unique segment. This implies that any boundary is included in one, and only one, of the two segments it bounds. The only possible quantitative evaluations between two events are: (i) whether they are in the same segment (*a* and *b* in fig. 1) or not, (ii) whether they are situated in contiguous segments (*a* and *c* in fig. 1), and if not, (iii) how many segments separate them (*a* and *d* in fig. 1). This computation of distances between segments in terms of segments is the essence of time reckoning. If two events occur within the same segment, it is impossible—as long as only one segmentation is involved—to situate them relative to each other and to evaluate the duration between their occurrences.

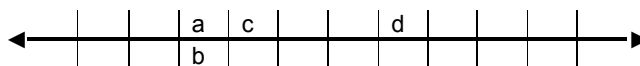


Fig. 1: Calendar division into identical segments; a-d mark events located within segments

Every calendar system employs a range of segmentations, which are inter-related in various ways, and used as a set of graduated rulers. Some segmentations produce quasi-equal periods that are clearly bounded from each other with no overlap, henceforth referred to as **UNITS** (e.g., days, months, years, in their strict calendar sense); other segmented periods have more vague boundaries, may exhibit overlap with consecutive segments, and vary in size relative to other segments of the

same cycle (most notably the seasons and the day segments, e.g., morning, noon, afternoon, evening, night, as well as day itself, in some of its uses). Thus, while (calendar) days are relatively equal units (depending on the calendar), clearly bounded and non-overlapping (such that each hour clearly belongs to a specific day), there is no crisp boundary between the periods that designate seasons or segments of the day, and neighbouring periods naturally overlap.<sup>5</sup> Segmentations of varying granularity are not necessarily synchronised.<sup>6</sup> Days are neatly contained within larger cyclic units without overlap (year, month, week, each a segmentation of relatively equal, clearly bounded, units), months (e.g., in the Gregorian calendar) are similarly contained within years, but the week is not aligned in the same way with either the monthly or the annual cycles.<sup>7</sup> Looser segmentations do not synchronise clearly with the larger calendrical unit containing them either—the period that reaches the boundary of the containing unit extends across it and into the next unit. In the Gregorian calendar, the night period extends across the dateline and winter extends beyond one calendrical year into the next.

The segmentation of the temporal line is abstract and conventionalised (that is, atemporal, based on non-temporal factors). Each segmentation forms its own sequence of consecutive segments, whether these are clearly bounded or overlapping periods (day following day, week following week, season following season, evening following afternoon and so on). In addition, each segmentation participates in (cyclic) sequences relative to larger granularities, which are composed of segments that are not directly consecutive on the same level (e.g., the first day of every month, the second week of every year, every Sunday, every morning). Formally, the temporal identification of any segment in any partition of a calendar system is determined by three components: the unit (type/size) of the temporal segment, the orientation (+, -) and the sequential number of the segment, starting with some central anchor-point. Thus, a temporal segment of type Day identified as +3 indicates the third segment in the time-unit Day from the anchor point.

Calendars involve two perspectives, one historical, relating segments to a primary anchor in the past, the other localised, anchored to the present. HISTORICAL CALENDARS are social calendars, and correspond to what is called absolute dating (depicted in fig. 2). In order to obtain historical significance, a calendar system needs a fundamental unit, typically a year, and an initial, establishing, anchor, a mythical or

---

<sup>5</sup> Referring here to calendar-internal designations and abstracting away from boundaries imposed through external means, such as legislation, religious prayer cycles, or work shifts.

<sup>6</sup> Some calendar segmentations are synchronised (to some level) with external phenomena (the rotation of the earth, the daily alternation of sunlight and darkness, the seasonal cycles, solar and lunar cycles, or rarer cycles such as the rotation of Venus or Sirius). There are, of course, discrepancies between the abstract calendar systems and the external phenomena at their basis, leading to a range of mechanisms intended to adjust and (re)align the various segmentations, through intercalation (the periodic addition of segments: leap seconds, days or months) or more directly through observation (most notably determining the beginning of a month through the observation of the new moon).

<sup>7</sup> This may be related to the fact that the week unit is not correlated with any natural cycle, unlike the other fundamental calendar units (see n. 6). By week, we refer to a named period, intermediate between days and months, but not necessarily equalling 7 days (weeks of 5, 6, 10, 13 and even 20 days have been used in various calendars over time).

historical event.<sup>8</sup> This anchor designates a boundary within a particular segmentation (E in fig. 2), and serves as the starting point for designating all other boundaries or segments. An historical calendar is composed of one or two parts, depending on the nature of the anchor event. E marks the initial boundary of the relevant segmentation, not precisely the anchor event itself. Thus, E properly belongs only to the sequence following it. Anchor events that allow a timeline preceding E, such as the Gregorian Calendar, have two parts, one starting with E and the segments following it, and one starting with E and preceding it. Thus, the timeline is segmented into two mirror-image ordinal sequences without a zero point between them, with E serving as a boundary rather than a segment in its own right. An anchor event that does not allow a preceding timeline (e.g., E designates the creation of the world as in the Jewish calendar) has a single, forward-looking sequence beginning with the E boundary (that is, only the right-hand side of fig. 2).

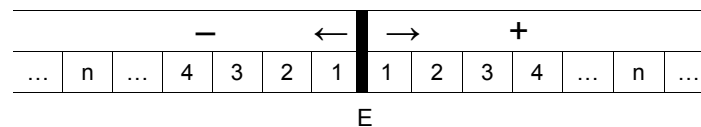


Fig. 2: The historical calendar system (two mirrored ordinal sequences without a zero)

This fundamental boundary is a fixed reference point from which all segments are computed. Ordinal numbers are the most relevant system for the denomination of the segments of a linear partition, since the entire timeline must be covered, that is, two infinite sequences must be named. This system enables the location of any event using three attributes: the location of the historical fundamental unit chosen (U0), the prospective or retroactive direction with respect to the location anchor (+, -), and its rank inside the chosen sequence. Other sequences of segments are either named through the same process (especially for larger units such as centuries) or using the fundamental numbering associated with a cyclic sequence of finite names (as the names of months or days in a week). Boundaries are not enumerated. Only segments need to be enumerated for dating purpose. By standard convention, the left boundary of a segment belongs to that segment, and its right boundary belongs to the next segment.

To deal with the notions of past, present and future, of temporal distance and proximity, localisation is done from a completely different perspective—tied to a variable anchor. It is this correlation that provides the temporal aspect of a calendar; the anchoring of perspective to a specific central segment (rather than to a boundary) which designates the present, the 'now', and the subsequent designation of two sequences of segments, forwards and backwards from this central point (past < present > future), is the necessary component that enables us to handle subjective,

<sup>8</sup> This event can be a concrete event (such as the beginning of a reign), but in its anchor role, it loses its own temporal duration and becomes a boundary—that is, it remains a part of the ground against which other events are temporally located, rather than being located itself. This is why there is no Year 0. The same was originally true for hours as well, which were enumerated using ordinal numbers, which exclude zero (except in Maya calendars, see Cauty, 2000). Nowadays, dates are often designated through cardinal numbers.



context-dependent aspects. This sequence of temporal periods enumerated from a central temporal anchor is the LOCALISED CALENDAR. This sequence is also potentially infinite in both directions. In the localised calendar, unlike in the historical calendar, the central segment is not included in the two sequences; the calendar is therefore composed of three parts (fig. 3). The same tripartite structure operates for both deictic and anaphoric uses. The only difference between the two is that in the deictic use anchoring is directly to the present, such that the central segment is the time of speech, while in the anaphoric use, anchoring is mediated and transposed, such that the central segment is the 'now' of a different point of view, a timeframe different from the time of speech.

—				←					→				+			
...	n	...	4	3	2	1	∅	1	2	3	4	...	n	...		

Fig. 3: The localised calendar system

Formally, the temporal identification of a deictic segment is also determined by three components: the size of the temporal segment, the orientation (future +, past -, present ∅), and the sequential number of the segment, starting with the central anchor (∅ in fig. 3). Thus, a temporal segment of type Day identified as +3 indicates a temporal segment three days in the future relative to the anchor segment ∅.

As noted above, linguistic reference to time imposes its own structures and patterns on this formal, abstract system. All three components of temporal identification (granularity, orientation, and sequence) are used in language, and all points in all cycles of the (culturally relevant) calendar can be referred to linguistically, of course; however, the necessary information is encoded differently in different grammatical constructions. The linguistic signal varies depending on the size of the segment selected, the intended orientation, and its proximity to the present anchor (∅), and according to various degrees of linguistic conventionalisation.

The study of the linguistic expressions associated with these deictic series reveals, for example, that S seems to never be translated to a number, nor are its two contiguous segments. The underlying formal structure of these linguistic deictic expressions exhibits a more complex relational and topological structure than the one offered by the set of integer numbers.

### 3. The Proximal Fundamental Units

Calendar units are linguistically referred to using lexicalised forms (*tomorrow*), frozen forms (*two months ago*), and free productive sequences (*a couple of weeks before Christmas*). The term 'frozen expression' is adopted here to refer to conventionalised strings that have lost productivity of form and temporal reference (*last year*, *the day before yesterday*). Frozen expressions are no longer freely variable, and often exhibit reduction of linguistic features that are typical of the relevant phrasal category, as

well as semantic standardisation of their temporal reference (even when temporal reference may still be compositionally retrievable). This aspect is illustrated in the English expressions referring to the deictic  $\pm 1$  units (*last year/month/week*, *next year/month/week*); these are clearly specialised for the  $\pm 1$  reference and lack a determiner, unlike other nominals modified by *last/next*. The equivalent productive strings which include determiners (*the last/next year*) are not interchangeable, and are not used as direct deictic reference in the same way. Such frozen expressions contrast with non-frozen deictic sequences, whose reference is not conventionalised, but remains compositionally computed (e.g., two days before yesterday).<sup>9</sup>

In language, the reference of the temporal deictic centre, the anchor, is identified as the time of speech (S), regardless of the size of the unit selected. Linguistic strings referring to the temporal deictic anchor itself are primary targets for conventionalisation. In many languages, linguistic reference to Day S is lexicalised, and reference to S in larger units (Week, Month, Year) is formed of the name of the unit modified by a determiner or a demonstrative. This type of modification can produce a complex or lexicalised expression, depending on properties of the specific language, as illustrated below.

Table 1: Expressions referring to the current S anchor of the fundamental units

	<b>S Day</b>	<b>S Year</b>	<b>S Month</b>	<b>S Week</b>
<b>English</b>	<i>today</i>	<i>this year</i>	<i>this month</i>	<i>this week</i>
<b>German</b>	<i>heute</i>	<i>dieses jahr</i> 'this.N year'	<i>diesen monat</i> 'this.M month'	<i>diese woche</i> 'this.F week'
<b>French</b>	<i>aujourd'hui</i>	<i>cette année (ci)</i> 'this.F year (this)'	<i>ce mois (ci)</i> 'this.M month (this)'	<i>cette semaine</i> 'this.F week '
<b>Estonian</b>	<i>täna</i>	<i>see aasta</i> 'this year'	<i>see kuul</i> 'this month'	<i>see nädal</i> 'this week'
<b>Polish</b>	<i>dzisiaj</i>	<i>tym roku</i> 'this year'	<i>tym miesiącu</i> 'this month'	<i>tym tygodniu</i> 'this week'
<b>Korean</b>	<i>onul</i>	<i>Olhae</i>	<i>īben-dal</i> 'this month'	<i>īben-ju</i> 'this week'

<sup>9</sup> We will not go into the processes that lead to this conventionalisation and whether any specific case should be regarded as grammaticalisation, lexicalisation, or some combination of the two. Both are diachronic processes of gradual language change in which linguistic strings lose free productivity and compositionality, becoming increasingly standardised in form and meaning. For an overview of the debate regarding the classification of these concepts and whether they are distinct processes or part of the same phenomenon, see Brinton and Traugott (2005) and references therein.

<b>Vietnamese</b>	<i>hôm nay</i> 'day this'	<i>năm nay</i> 'year this'	<i>tháng này</i> 'month this'	<i>tuần này</i> 'week this'
<b>Arabic<sup>a</sup> (Lebanese)</b>	<i>al ywm</i> 'DF day'	<i>as senh</i> 'DF year'	<i>aš šahr</i> 'DF month'	<i>al usbu</i> 'DF week'
<b>Hebrew</b>	<i>ha.yom</i> 'DF.day'	<i>ha.šana</i> 'DF.year'	<i>ha.xodeš</i> 'DF.month'	<i>ha.šavua</i> 'DF.week'

The formal three-part process of identification for non-S references is evident in constructions where all three elements—type of unit, orientation, and sequential number—are explicitly and independently identifiable (see Tables 2–3). Such strings are available from a certain number of units (e.g., two and higher, depending on granularity and language), are not (linguistically) upper-bound, and are typically the only option available for distal reference. However, for more proximal units ( $\pm[1-3]$ , depending on language), this transparent compositional option is not the only, or the preferred, option. Rather, in this proximal range, speakers tend to prefer lexicalised or frozen strings in which the computational information is opaque to varying degrees, and temporal reference is not entirely compositional. This is illustrated by an expression such as the lexicalised French *après-demain*, which refers directly to the +2 day following Day S, although literally indicating an open-ended time period that begins after tomorrow, a term which is itself opaque. In fact, the parallel non-lexicalised string (*après demain* lit. 'after tomorrow') is also in use, with the expected open-ended temporal reference. Typically, the conventionalised strings refer directly to the intended temporal segment without recourse to additional temporal information (such as verbal time reference).<sup>10</sup> Note that the existence of conventionalised constructions does not preclude the use of the compositional numerical string (e.g., Eng. *in n days*) as well. All the languages exemplified in Tables 2–3 allow the numerical construction starting from  $D\pm 2$  and  $Y\pm 2$ .

<sup>a</sup> The allomorphic changes to the DF marker in Arabic are a result of phonological assimilation (see Watson, 2002: 216–222).

<sup>10</sup> In some cases, the information in the conventionalised expression is not sufficient to determine reference without additional information. This is illustrated by languages in which the same form serves both orientations (D+1 and D-1: *kal* in Hindi and Urdu, *ejo* in Kinyarwanda (Bantu), D+2 and D-2, *parsō* in Hindi and Urdu, *paramdivas* in Gujarati). These words can be further modified to show orientation. Other units in these languages do show regular marking for orientation. These languages have not been examined in the present study.

Table 2.1: Expressions referring to the proximal sequence of days

	English	French	Polish	Estonian
D+n	<i>in n days</i>	<i>dans n jours</i> 'in n day.PL'	<i>za n dni</i> 'in n day.PL'	<i>n päeva pärast</i> 'n day.PL after'
D+3	<i>in 3 days</i>	<i>après-après-demain</i> 'after.after.tomorrow'	<i>popojutrze</i> 'after.after.tomorrow'	<i>üle-ülehomme</i> 'beyond.beyond.tomorrow'
D+2	<i>the day after tomorrow</i>	<i>après-demain</i> 'after.tomorrow'	<i>pojutrze</i> 'after.tomorrow'	<i>ülehomme</i> 'beyond.tomorrow'
D+1	<i>tomorrow</i>	<i>demain</i>	<i>Jutro</i>	<i>Homme</i>
D0	<i>today</i>	<i>aujourd'hui</i>	<i>Dzisaj</i>	<i>Täna</i>
D-1	<i>yesterday</i>	<i>hier</i>	<i>Wczorai</i>	<i>Eile</i>
D-2	<i>the day before yesterday</i>	<i>avant-hier</i> 'before.yesterday'	<i>przedwczoraj</i> 'before.yesterday'	<i>üleeile</i> 'beyond.yesterday'
D-3	<i>3 days ago</i>	<i>avant-avant-hier</i> 'before.before.yesterday'	<i>przed-przedwcoraj</i> 'before.before.yesterday'	<i>üle-üleeile</i> 'beyond.beyond.yesterday'
D-n	<i>n days ago</i>	<i>il y a n jours</i> 'EM n days'	<i>n dni temu</i> 'n day.PL ago'	<i>n päeva tagasi</i> 'n day.PL ago'

Table 2.2: Expressions referring to the proximal sequence of days

	Vietnamese <sup>b</sup>	Hebrew	Malagasy
D+n	<i>trong n ngày nữa</i> 'in n day more'	<i>od n yamim</i> 'in n day.PL'	<i>afaka (n-1) andro</i> 'go-over (n-1) day'
D+3	<i>ngày kia</i> 'day a-bit-further'	<i>od 3 yamim</i> 'In 3 day.PL'	<i>afaka 2 andro</i> 'go-over 2 day'
D+2	<i>ngày kia</i> 'day a-bit-further'	<i>Maxratayim</i> 'tomorrow.DUAL-PL'	<i>raha afaka rahampitso</i> 'If-manage go-over tomorrow'
D+1	<i>ngày mai</i> 'day after'	<i>Maxar</i>	<i>rahampitso</i>

<sup>b</sup> Marking of gradually increasing distance is achieved through tone alternations on *kia*. This construction can be used up to  $\pm 8$  in poetic registers, but is rarer above  $\pm 3$  in spoken registers (Xuyen Lê Thi, PC).

D0	<i>hôm nay</i> 'day this'	<i>ha.yom</i>	<i>androany</i>
D-1	<i>hôm qua</i> 'day passed'	<i>Etmol</i>	<i>omaly</i>
D-2	<i>hôm kia</i> 'day a-bit-further'	<i>Šilšom</i>	<i>1 andro lasa</i> '1 day done'
D-3	<i>hôm kia</i> 'day a-bit-further'	<i>lifney 3 yamim</i> 'before 3 day.PL'	<i>2 andro lasa</i> '2 day done'
D-n	<i>cách đây n hôm</i> 'distance here n day'	<i>lifney n yamim</i> 'before n day.PL'	<i>-(n-1) andro lasa</i> '-(n-1) day done'

Table 3: Expressions referring to the proximal sequence of years

	German	Estonian	Arabic (Algerian)	Vietnamese
Y+n	<i>in n Jahren</i> 'in n year.PL'	<i>n aastal pärast</i> 'n year after'	<i>mena aala n snine</i> 'from now n year.PL'	<i>trong n năm tới</i> 'in n day coming'
Y+3	<i>in 3 Jahren</i> 'in 3 year.PL'	<i>3 aastal pärast</i> '3 year after'	<i>mena aala 3 snine</i> 'from now 3 year.PL'	<i>trong 3 năm tới</i> 'in 3 year coming'
Y+2	<i>in 2 Jahren</i> 'in 2 year.PL'	<i>üle-tulval aastal</i> 'beyond.coming year'	<i>mena aala 3aamine</i> 'from now year.DUAL-PL'	<i>trong 2 năm tới</i> 'in 2 year coming'
Y+1	<i>nächstes Jahr</i> 'next year'	<i>tulval aastal</i> 'coming year'	<i>el 3aam el jay</i> 'DF year DF coming'	<i>năm tới</i> 'year coming'
Y0	<i>dieses Jahr</i> 'This.M year'	<i>see aasta</i> 'this year'	<i>had el 3aam</i> 'this DF year'	<i>năm nay</i> 'year this'
Y-1	<i>voriges Jahr</i> 'previous year'	<i>eelmisel aastal</i> 'previous year'	<i>el 3aam li faat</i> 'DF year that passed'	<i>năm ngoái</i> 'year passed'
Y-2	<i>vor 2 Jahren</i> 'before 2 year.PL'	<i>üle-eelmisel aastal</i> 'beyond.revious year'	<i>hadi 3aamine</i> 'that year.DUAL-PL'	<i>cách đây 2 năm</i> 'distance here 2 year'
Y-3	<i>vor 3 Jahren</i> 'before 3 year.PL'	<i>3 aastat tagasi</i> '3 year ago'	<i>hadi 3 snine</i> 'that 3 years'	<i>cách đây 3 năm</i> 'distance here 3 year'
Y-n	<i>vor n Jahren</i> 'before n year.PL'	<i>n aastat tagasi</i> 'n year ago'	<i>hadi n snine</i> 'that n years'	<i>cách đây n năm</i> 'distance here n year'

Our study indicates that temporal proximity is linguistically expressed along two dimensions: sequential and scalar. The sequential dimension is related to the

distance of a unit relative to the anchor of the same unit, and reflected in the changes in linguistic structures associated with the various segments in the sequence. The languages examined demonstrate a correlation between the level of temporal proximity to the deictic anchor and the diachronic process of conventionalisation, whether to a lexical level or to a larger frozen string. The scalar dimension is related to granularity, to the type of unit, and reflected in the wider expanse of the proximal range in the Day unit compared to the Year unit (cf., Tables 2–3; note that in some languages (e.g., Estonian, French, Vietnamese) the proximal Day constructions can be extended beyond  $D\pm 3$ , but these forms are less-frequently used). Day S and its immediate  $\pm 1$  neighbours are most frequently lexicalised, encapsulating their standardised temporal reference opaquely, such that the components of temporal identification are, typically, not compositionally accessible to the speakers. The more proximal units in both dimensions provide a richer vocabulary beyond the numerical strings, and the closest proximal units are not designated by numerical strings at all (cf., Smith, 2003). Based on the range of constructions employed in the languages examined, and the clear patterns of conventionalisation, we propose that the anchor of the present (S) conceptually determines three concentric zones in the temporal domain—the central, the proximal, and the distal—which extend both to the future and the past, and are separated by fuzzy boundaries (fig. 4).

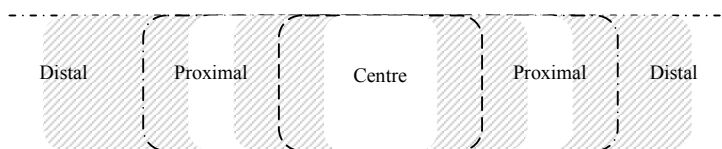


Fig. 4: Zones of deictic temporal reference

Linguistic reference to times in the distal zone employs the compositional constructions of sequential reference (e.g.,  $D\pm n$  in Tables 2–3), and reference to the proximal and central zones employs specialised constructions. While this zoning is found in all languages examined, and encoded in the relevant linguistic forms, it is realised through various linguistic means in different languages. The inventory of linguistic methods evident in our sample languages includes the use of determiners, adjectives, and prepositions (evident in all examined languages), morphological affixation (French, Estonian, Polish, Hebrew), reiteration (Estonian, Polish), as well as tonality and prosody (Vietnamese).<sup>11</sup>

Lexicalisation is typically limited to the 0–1 range, although some languages exhibit a somewhat wider range. The effect of the type of unit is very noticeable here; the day unit most frequently takes the most conventionalised, lexicalised forms, that are the most opaque. This suggests that the day unit is more primary; it appears to be the most frequently used unit, which may account for its higher level of

<sup>11</sup> This range nicely illustrates Levinson's (2005: 112) generalisation that although deictic categories are universally reflected in language, no universality can be assumed regarding the specific linguistic realisations of these categories.

conventionalisation. The year unit exhibits more conventionalised constructions compared to the week and month.

The distal zone is only bounded on its proximal periphery. This is evident in the absence of a linguistic upper bound to the numerical temporal constructions. Numerical compositional structures are available from  $\geq 3$  (in some languages even starting with 2), and can be used alongside the specialised constructions of the proximal zone, providing speakers with more choice. Given the multiple forms of expression, the size of the intersection (and overlap) between the numerical and non-numerical forms varies not only from language to language, but also among speakers. Further empirical study is required to examine to what extent speakers employ specialised forms for more distant temporal segments, or whether these forms are going out of use. The limited use of existing conventionalised strings that refer to more distant segments is evident in the reiterated affixation of Estonian *üle* 'beyond' or French *après* 'after' and *avant* 'before' (see Table 1). Although productive cyclic reiteration indicating progressive distance of days is morphologically possible in both languages and judged grammatical, speakers show a clear preference to employ the numerical construction for more distal periods, and affixation of more than 2–3 duplicates is rare. The more distant reiterated forms are not standardised and their temporal reference requires computation, a fact which correlates and is probably related to their limited use. A similar situation is found with the phonological marking of increasing distance in Vietnamese (see n. b, Table 2.2). Note, in the same context, the Hebrew lexicalised form *eštakad* 'last year', which is now rarely used and limited to higher registers.

The languages examined reveal that the three zones vary in size across languages and across units. The Day unit, which is much more frequently opaquely lexicalised, also shows a wider extension of the proximal and central zones compared to the Year and the intermediate fundamental units. In addition, the zones are not necessarily symmetric in the future/past orientation. Asymmetry is reflected in languages where the central and proximal zones extend further in one orientation or another. This is illustrated in Albanian, Farsi and Russian, where conventionalisation is evident further to the past than to the future, for some units. This asymmetry is not surprising as it hinges on conventionalisation processes that are non-uniform and idiosyncratic to a certain extent, particularly in stages of lexicalisation.

As noted above, the distributional properties of these expressions in the languages examined indicate a nominal categorisation, although their functions in the clause may be adverbial (as evident in their traditional categorisation in some languages). We therefore consider them nominal and referential expressions on par with other nominal forms. Consequently, the zoning pattern can be correlated with the scale of referential activation/ accessibility/ givenness that is associated with referring expressions in pragmatic theory (cf., Chafe, 1976; Ariel, 1990; Gundel et al., 1993).<sup>12</sup> Accordingly, more opaque (and more central) forms refer to more activated, relatively accessible referents in the addressee's mind, while the very transparent numerical structures refer to temporal referents of low accessibility. Thus, the

---

<sup>12</sup> Functional pragmatic theories of referential expressions argue that the form of referring expressions linguistically encode or signal the discourse functions/ cognitive statuses that the intended referent is assumed to have in the mind of the addressee, thereby constraining possible interpretations. (e.g., pronouns are 'active' in Chafe, 1976, 'highly accessible' in Ariel, 1990; 'in focus' in Gundel et al., 1993, while new referents are 'inactive', 'low accessible', and 'type identifiable' respectively).

proximal and central Day units are also the highest in accessibility among these forms.

#### 4. The Day Segments

The segments of the Day (such as morning, noon, afternoon, evening, night) exemplify a different type of segmentation. This partition is composed of non-equivalent, sequential but overlapping segments, whose boundaries are vague. This type of segmentation is not in a part-whole relation with the fundamental calendar units which it segments. In this case, day segments are smaller than a calendar day, but can overextend the boundaries between calendar days. Day segments can be used to refer to two series, one within the containing calendar Day unit (*this morning* vs. *this afternoon*), and the other by type (*this morning* vs. other mornings). As with the fundamental units, comparative examination of the linguistic use of this segmentation reveals recurring patterns. The languages examined were English, French and Hebrew.

The absence of crisp boundaries is significant for the linguistic use of these expressions, since boundaries may vary with each token. Consequently, the same point in time can belong to different segments, even within the same discourse, as illustrated in the French example in (1).

- (1) [...] depuis le mardi dès le soleil couché fixé à cinq heures du soir, relativement au règlement des eaux du 22 juin 1748, jusqu'au mercredi à **deux heures après midi** et depuis le samedi à onze heures du matin jusqu'au dimanche à **deux heures du soir**, ce qui compose 48 heures dans chaque semaine pour servir depuis le 1<sup>er</sup> mars jusques après le détritage des olives du moulin du dit Cirlot et de la Communauté. ([http://www.pays-du-var-est.eu/le\\_ray-23.htm](http://www.pays-du-var-est.eu/le_ray-23.htm))

'[...] from Tuesday at sunset, fixed at 5 o'clock in the evening, following the water regulation of June 22, 1748, until Wednesday **at 2 o'clock in the afternoon** and from Saturday at 11 o'clock in the morning until Sunday **at 2 o'clock in the evening**, which make up 48 hours every week, to serve from March 1st until after the pressing of the olives of the Cirlot mill and of the community.'

As with the fundamental units, proximity to the deictic centre affects the linguistic use: central and proximal reference showing more conventionalised constructions, with numerical constructions available from  $\pm 2$ . Lexicalised and frozen expressions used with day segments are not necessarily parallel to those used for the fundamental units; conventionalisation may be idiosyncratic among the various day segments, with some periods exhibiting more frozen, opaque constructions than others. English marks segments of Day 0 using two central patterns, lexicalised (e.g., *tonight*) and a frozen demonstrative (e.g., *this morning/evening*), both of which are operative with the fundamental units. French and Hebrew show a more uniform paradigm, paralleling the marking of the fundamental units in those languages: the segments of Day 0 are marked with a demonstrative in French and a definite marker in Hebrew. Marking of the  $D \pm 1$  segments differs from that of the fundamental units and variant within the different day segments in all three languages (Eng. *\*last/next*



*morning* vs. *last night/evening* \**next night/evening*; Fr. \**le matin/soir dernier/prochain* vs. *la nuit dernière/prochaine*) (on French, Berthonneau, 1989). The various periods within this segmentation differ in status, as reflected by the conventionalised constructions referring to them. The more central segments of the day, such as *morning*, *evening* and *night* are commonly used with central and proximal markings, while less prominent segments are less systematic: reference to noon (Eng. *noon* / Fr. *midi* / Heb. *caharayim*), a brief, in some cases (almost) punctual segment, is not marked for deictic proximity; reference to the afternoon segment (Eng. *afternoon* / Fr. *après-midi* / Heb. *axrei-ha.caharayim*) shows S marking, but is less common with conventionalised  $\pm 1$  forms.

Unlike the fundamental units, the day segments in our sample languages cannot provide independent temporal identification, and require additional linguistic cues, such as tense, to identify the intended segment. Thus, Eng. *this afternoon* and Fr. *cet après-midi* can refer to a present, past or future orientation, cued by the time of S during the present day and by the choice of tense.

- (2) S=10:00, future markings on the predicate
  - a. This afternoon, I will pick Claire up from school.
  - b. *Cet après-midi, j'irai chercher Claire à la sortie de l'école.*  
'This afternoon, I will look for Claire at the school gate.'
  - c. *axrei ha.caharayim, ani ekax et Claire mi.bet-sefer.*  
'This afternoon, I will take Claire from school'
- (3) S=15:15, 18:15, present markings on the predicate<sup>13</sup>
  - a. This afternoon, we are working on calendars.
  - b. *Cet après-midi, nous travaillons sur les calendriers.*
- (4) S=21:00, past markings on the predicate
  - a. This afternoon, we had very interesting presentations.
  - b. *Cet après-midi, nous avons eu des exposés très intéressants.*  
'This afternoon, we had very interesting presentations.'
  - c. *axrei ha.caharayim hayu harca'ot mamaš me'anyenot.*  
'This afternoon, there were very interesting lectures'.

This lack of independent reference is particularly evident with respect to the night segment. Words referring to day and night involve an alternation which does not

<sup>13</sup> The present '*benoni*' in Modern Hebrew gets a habitual or a *future proche* reading unless explicitly marked with *axšav* '(right) now' or a similar modifier. This constrains the use of the temporal markers under discussion. This aspect clearly requires further examination, but does not detract from our argument here, that day segments are not specified for orientation and the missing temporal information is completed through other means.

correspond to calendar dates (in the Gregorian calendar), but is the mundane light/dark division in which night overlaps the calendar Day unit. A calendar day includes two successive but disjoint night segments, one preceding the dawn and another following sunset. In the three languages examined, the relevant linguistic expressions (Table 4) refer to either one of the two non-adjacent night periods, preceding and following the relevant day, regardless of the calendar D unit, as illustrated in (5–6).

- (5) a. We'll go out tonight. (S=8:00; 17:00)  
 b. There'll be a meteor shower tonight at 2 a.m. (S=anytime from morning)  
 c. We have a really special dessert tonight. (S=during dinner)  
 d. We had to go to the ER tonight. (S=22:00; 4:00; ??8:00)  
 e. Tsunami waves hit Indonesia tonight at 18:34. (S=22:00; 4:00; ?8:00)

- (6) *La nuit d'hier a été chaude du côté des halles Saint-Martin, où la police est intervenue à deux reprises, en peu de temps. Il était 23h, avant-hier, lorsqu'un homme de 56 ans a voulu retirer de l'argent au distributeur fixé dans le mur des halles. [...] Puis, vers 0h40, place Maurice-Gilet, au pied de l'église [...]*

'**Last night** was heated at the Saint-Martin hall, where the police intervened twice in a short time. It was **23:00 the day before yesterday**, when a man of 56 wanted to withdraw money from the ATM in the wall of the hall [...] Then, around **0:40 AM**, at Maurice-Gilet square at the foot of the church [...]

(<http://brest.letelegramme.com/local/finistere-nord/brest/ville/double-agression-dans-la-nuit-a-saint-martin-02-07-2011-1357318.php>)

Within a specific utterance, the intended reference is generally non-ambiguous, thanks to the linguistic (and contextual) cues that supply the missing orientation, as illustrated in (5–6). Appropriateness of a specific orientation is strongly affected by the exact time of S within the Day 0 anchor, although there is much speaker variability with respect to the specific boundaries. Thus, speakers are more likely to use the 0-marked form (*tonight*) to refer to the preceding part of the night if S is either prior to midnight or prior to sleep, but choose the -1-marked form if S is after midnight or sleep.<sup>14</sup> The significance of these situational factors is highlighted in the fact that, out of context, some English speakers accept only the possibility of a future reference with *tonight*, a perception clearly contrasted by attested data in context, even by those speakers (as in 5).

<sup>14</sup> Similarly, speakers working a night shift choose the 0-marked form during the shift but are more likely to use the -1-marked form once the shift is over.

Table 4: Expressions referring to proximal nights (English, French, Hebrew)

	-2	-1	0	+1	+2
	D	D	D <sub>s</sub>	D	D
	N-2,-1	N-1,0	N0,1	N1,2	
Eng.	<i>last night</i> * <i>yesterday night</i>	<i>tonight</i> <i>last night</i>	<i>tonight</i> * <i>next night</i>	<i>tomorrow night</i> * <i>next night</i>	
Fr.	<i>la nuit derniere</i> 'DF night last' <i>la nuit d'hier</i> 'DF night of. Yesterday'	<i>cette nuit</i> 'this night' <i>la nuit derniere</i> 'DF night last' <i>la nuit d'hier</i> 'DF night of. Yesterday'	<i>cette nuit</i> 'this night' <i>la nuit prochaine</i> 'DF night next'	<i>la nuit prochaine</i> 'DF night next' <i>la nuit de demain</i> 'DF night of tomorrow'	
Heb.	<i>etmol ba.layla</i> 'yesterday in. DF.night'	<i>ha.layla</i> 'DF.night' <i>etmol ba.layla</i> 'yesterday in.DF.night'	<i>ha.layla</i> 'DF.night' <i>maxar ba.layla</i> 'tomorrow in. DF.night'	<i>maxar ba.layla</i> 'tomorrow in. DF.night'	

This segmentation of the day includes a wide range of segments that vary in properties and usage (among which, aside from all segments already mentioned above, are dawn, twilight, sunrise, midday, dusk, sunset, midnight). Two points are crucial for the interpretation of this segmentation language: (i) the dependence of these expressions on additional cues to determine orientation, and therefore exact temporal reference; (ii) appropriateness of reference does not depend entirely on distance from S, but on fuzzy temporal boundaries that are neither linguistic nor purely calendrical, such as sleep or work shifts. This correlates with the central features of this segmentation—the vaguely defined and overlapping boundaries.

## 5. Named Days of the Week

The third type of expression, the named days of the week, is an aggregation of consecutive occurrences. Formally, the names of days (e.g., *Friday*, *vendredi*) define a disjoint sequence, each element of which is associated with a unique occurrence within the unit of the next granularity, the week.

The part-whole relation between a named day and a week is a complete calendar part-whole relation, but linguistically, this relation is less evident. Rather, temporal reference depends on the proximity of Day 0 rather than on inclusion within the week segmentation (with orientation necessarily provided by the verbal temporal reference as it is with the segments of the day). This is exemplified in temporal reference with unmodified named days in French. The use of a bare named day identifies the nearest occurrence of a day of this type relative to S (excluding the type of Day S itself). So, if S is on Monday 0, the use of *mardi* can refer to Tuesday 0 or Tuesday -1, the use of *jeudi* to Thursday 0 or Thursday -1, but *lundi* cannot refer to Monday 0, nor to Monday  $\pm 1$ . The same effect is evident with days of the week

introduced by a demonstrative, another 0-marking strategy with named days of the week in French. Thus, the temporal identification of *ce jeudi* 'this Thursday' depends on the proximity of Day 0 rather than on its inclusion in Week 0, with past or future orientation provided by additional elements in the context.

In contrast, named days marked for lower proximity in French explicitly encode the orientation of reference, specifically through modification by *dernier* / *prochain* 'last / next'. Despite this additional information, the exact identification of the intended temporal referent is still ambiguous. The data provided by our informants suggest that these constructions remain unclear as to whether the intended reference is the nearest occurrence or the second in the sequence (similar effects have also been reported by informants in English and Hebrew). Although the resulting ambiguity is affected by the relative proximity of Day 0, it is not resolved in specific occurrences. For the majority of speakers, interpretation is based on proximity to Day S, excluding reference to the central zone, which is too close. The further the day-type mentioned is from Day S, the more people interpret the Week 0 occurrence of that day as the intended referent. Thus, if S is on Monday 0, the four day gap to coming Friday of the same week is sufficient distance to assume that this is the intended referent of *vendredi prochain* 'next Friday', while Tuesday, the immediately following day, is too close to be the intended referent of *mardi prochain* 'next Tuesday'. In the past orientation, if S is on Friday 0, the gap to the preceding Monday of the same week provides sufficient distance to assume that this is the intended referent of *lundi dernier* 'last Monday', while the immediately preceding Thursday is too close to be the intended referent of *jeudi dernier* 'last Thursday'. This combination of factors produces the following gradient judgements (calculated for a homogenous sample group of over 10 people).<sup>15</sup>

Table 5: Temporal reference judgements on proximal future named days (French)

S= Monday 0	Named Day in Week 0	Ambiguous	Named Day in Week +1
<i>mardi prochain</i> 'Tuesday next'	--	--	<b>100%</b>
<i>mercredi prochain</i> 'Wednesday next'	16%	8%	<b>75%</b>
<i>jeudi prochain</i> 'Thursday next'	25%	<b>50%</b>	25%
<i>vendredi prochain</i> 'Friday next'	<b>75%</b>	16%	8%

<sup>15</sup> For expository clarity, reference is presented as anchored to the including week, as this is the only anchoring that remains constant with the changing S and target strings.

Table 6: Temporal reference judgements on proximal past named days (French)

S= Friday 0	Named Day in Week 0	Ambiguous	Named Day in Week -1
<i>jeudi dernier</i> 'Thursday last'	--	--	<b>100%</b>
<i>mercredi dernier</i> 'Wednesday last'	16%	8%	<b>75%</b>
<i>mardi dernier</i> 'Tuesday last'	25%	<b>50%</b>	25%
<i>lundi dernier</i> 'Monday last'	<b>75%</b>	16%	8%

The distance from Day S evidently has a crucial impact on the identification of the intended referent for most speakers. The variations of reference with these expressions can be attributed to two factors: the fuzzy boundary of the zones in general, already discussed above, and more specifically, the nature of this particular segmentation as a sequence of Days with a strict correlation to the Week unit (at least its basic definition). The latter factor leads speakers to question whether reference should be obtained through the sequence of days, in which case reference is to the nearest occurrence of the named day, or whether it should be obtained through the sequence of weeks, in which case, the intended referent is the one in the Week  $\pm 1$ . For a small set of speakers, interpretation is based entirely on the Week, leading them to consistently identify temporal reference in Week  $\pm 1$  regardless of distance from Day S.

To conclude this section, successful temporal reference with this segmentation involves factors such as the distance from S and the encoding (or absence) of orientation in the expression itself, which have been shown above to play a significant role with other segmentations. In addition, the named days of the week indicate that ambiguity of reference may be present (between interlocutors, but crucially also for the same interlocutor) even when orientation and distance are encoded in the linguistic expression, and not clearly resolved even in specific contexts. Speaker awareness of this potential ambiguity leads to the addition of more temporal markers in the hope to clarify reference, increased specification which may be interpreted as redundant. We leave this topic to a future study.

## 6. Typological Indexing from a Formal Perspective

In this section, we would like to revisit the underlying systems depicted by deictic calendar expressions. The representation provided in section 2 is compatible with a representation of physical time in which any past event precedes any current event, and any current event precedes any future event (i.e. compatible with the relative temporal B-series of McTaggart, 1908). This order puts these series in a one-to-one correspondence of any deictic calendar segmentation with the set of integer numbers

Z (fig. 5), translating S to 0. As noted in our introduction, this correspondence is used extensively in computational linguistics and natural language processing.



Fig. 5: The Z-order

However, there is a significant difference, in mathematical terms, between the Z type of order and the tripartite structure of the localised calendar (fig. 3). The Z type order is a single bi-infinite sequence (defined by the relation  $x$  precedes  $y$  if  $y-x > 0$  [thus, -3 precedes 2 because  $2-(-3)=5$ ]) and not two separate infinite sequences, one for positive numbers and one for negative ones. In contrast, the tripartite structure is composed of three linear orders, one finite order restricted to a single element (denoted 0) and two infinite sequences starting with a minimal element (denoted 1), that is, two ordinal sequences linked through a central articulation. As indicated by the discussion above, linguistic use of calendar expressions is based on the latter, tripartite structure. Furthermore, the two infinite sequences, which appear to be mirror-images in mathematical terms, are revealed as non-parallel linguistically. In addition, linguistic use reveals the significance of the notion of distance, and of subjectivity of distance, which are not taken into account in a mathematical perspective. In light of the linguistic patterns discussed above, we can now propose a more cognitively and linguistically motivated representation of the deictic calendar segmentations.

### 6.1 The Underlying Order

First, linguistically speaking, deictic reference to temporal segments does not employ a single super-sequence. Rather, it appears that any linguistic sequence of deictic calendar expressions is divided into three parts: a central, finite section, which does not use transparent compositional numerical strings, and two potentially infinite sequences, which employ the transparent numerical strings. The infinite sequences do not necessarily begin with 1, and do not necessarily start from the same number (cf., the expressions for  $U \pm 2-3$  in Tables 2–3). Multiple expressions of various types may be used for the same temporal referent, numerical and non-numerical, transparent and opaque, freely productive or frozen (e.g., *in 2 days* vs. *the day after tomorrow*).

Second, it has been shown that the two infinite sequences are not necessarily symmetric, either in their starting point or in the conventionalised linguistic structures employed in the two orientations. Such asymmetry between orientations is evident in temporal expressions associated with a spatial orientation (vertically, horizontally or in depth). In Malagasy, past orientations employ the horizontal axis, while future oriented expressions employ the vertical axis, specifically represented as going over a physical obstacle, reminiscent of show jumping. Similar verticality is evident in German deictic expressions (*übermorgen* 'the day after tomorrow' lit. 'over.tomorrow') and in French anaphoric expressions, where verticality applies in both orientations (*surlendemain* '2 days later' lit. 'over.following-day' and *surveille* '2 days before' lit. 'over.preceding-day', although another non-vertical lexicalisation is

preferred for the latter reference). Another example of asymmetry between orientations is evident in the use of sequential markers of proximity illustrated by the non-opposing markers *next U* and *last U* (but *\*previous U* and *\*first U*).

The use of the ordinal marker *last* (and equivalents cross-linguistically) in deictic temporal constructions<sup>16</sup> is suggestive of the underlying structure involved. In non-temporal contexts, the use of this marker theoretically (i) indicates a finite sequence, (ii) refers to the only element in this finite sequence without a successor, (iii) is supposed to be associated with the ordinal adjective *first*, which refers to the only element of the sequence without a predecessor, (iv) resolves any need to know the cardinality of the sequence in order to refer to the last element, to the preceding element, if one exists, and so on. Within temporal deictic expressions, this ordinal marker (i) appears in a potentially infinite sequence, (ii) refers to the contiguous past segment of the anchor segment, and therefore is in natural correspondence with the number 1, (iii) is not associated with parallel deictic expression marked by *first*. The immediate neighbour of the anchor to the past is both the previous anchor and the first segment in this orientation, irrespective of its depth; the past sequence therefore functions as a Last In First Out stack structure. The future sequence can be viewed as a First In First Out queue structure, as the immediate neighbour of the anchor to the future is the first segment in the sequence and also the next anchor.

Finally, it has been shown that distance plays a crucial role in the conventionalisation of linguistic structures and in their synchronic use. It should be stressed that distance effects are not determined solely on the basis of the strict temporal sequence. The subjectivity of distance and the communicative significance of the expressions chosen in signalling subjective importance or urgency of the dated event are beyond the scope of the present work, but clearly impact on the linguistic use of these expressions.

These observations lead us to reject the assumption that the underlying order of a deictic calendar series is in a natural primary correspondence with the  $\mathbb{Z}$  set of integers (fig. 5), since its elements do not form the expected prototypical, discrete, and unbounded linear order. Rather, we propose a more complex representation intended to capture the linguistic complexity. This order is the result of an inference reasoning on the composition of two levels of ordering. A qualitative temporal level, corresponding to the well-known past–present–future order, or the more fine-tuned partition—remote past–proximal past–present–proximal future–remote future—that is evident in (some) tense systems, followed by a quantitative order within each section. The quantitative order is linear, either finite or non-finite with at least one item. The three more proximal sections (the central and proximal zones of fig. 4) are finite. Both remote past and remote future are unbounded and, independently of each other, in a one-to-one correspondence with the ordinal numbers from the relevant starting point (most commonly  $\pm 2$  or 3). Figure 6 illustrates the basic tripartite structure, with infinite remote past and future sequences and a finite central section. The most accessible segment in the past sequence is the segment last to have been in the centre. Consequently, the past sequence is viewed as an infinite stack structure (Last In First Out), with the most accessible segment on top. The most accessible segment in the

<sup>16</sup> Contrary to the historical or anaphoric calendar, in which linguistic expressions can include explicit ordinals (Eng. *the third year*, Fr. *le troisième jour*) or cardinals with an ordinal meaning (Eng. *In 2011*, Fr. *En 2011* for the 2011<sup>th</sup> year after Christ), deictic linguistic expressions typically do not employ ordinals. Deictic numerical constructions employ cardinals to signal the distance to the intended referent.

future sequence is the segment next to be in the centre. Consequently, this sequence is viewed as an infinite queue structure (First In First Out), where the most accessible segment is the head of the queue.

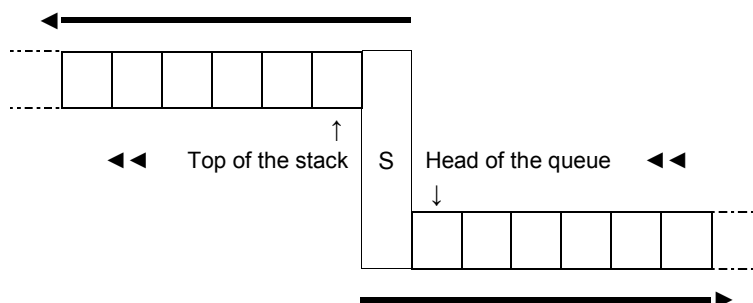


Fig. 6: The linguistics deictic calendar order type

This complex structure corresponds to Aristotle's definition of the relation prior-in-time with respect to events (*Metaphysics*, Book V, part 11):

"Other things are prior in time; some by being farther from the present, i.e. in the case of past events (for the Trojan war is prior to the Persian, because it is farther from the present), others by being nearer the present, i.e. in the case of future events (for the Nemean games are prior to the Pythian, if we treat the present as beginning and first point, because they are nearer the present)."

For Aristotle, there are three series of events: the series of past events, which is structured by the order prior that equals farther, the series of future events, which is structured by the order prior that equals nearer, and the series of present events. There is no direct continuity between a past event and a future one; comparison is only inferred through the qualitative relation, such as past prior to present prior to future. During this inference process, one of the two infinite orders is reversed.

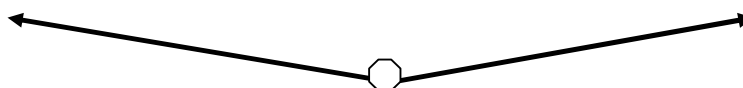


Fig. 7: Aristotle's prior order



## 6.2. Labelling and Enumeration of Reference Points

In theory, reference points along the two deictic orientations can be expressed through:

- numerical strings starting with 2 (line 1 in Table 7)
- conventionalised strings (lexicalised or frozen) that indicate the alternation of reference by single steps without the use of numbers (line 2).

Table 7: Labelling of temporal reference points (sequential)

– ←						→ +				
nU-	...	3U-	2U-	U-	S	U+	2U+	3U+	...	nU+
	...	U---	U--	U-	S	U+	U++	U+++	...	

The numerical strings are typically transparent and compositional, including explicit unit type, orientation and number in the sequence (e.g., 3U- appearing as *3 days ago*). The theoretical options in line 2 can be linguistically realised in various ways, given the range of lexical and non-lexical strings possible. In the following patterns, U represents the relevant unit, square brackets represent word boundaries.

- S and  $U_{\pm}$  can be realised as lexicalised  $[U_S]$  and  $[U_{\pm}]$  or as modified strings  $[U]_{\pm}$ . All these options are commonly attested cross-linguistically.  $[U_{\pm}]$  forms are the most frequent basis for forms of further proximity.
- $U_{\pm\pm}$  can be lexicalised as either  $[U_{\pm\pm}]$  or  $U[\pm\pm]$ , or as a modified form  $[U_{\pm}]_{\pm}$  on the basis of a prior lexicalisation  $[U_{\pm}]$ . All options are attested.
- $U_{\pm\pm\pm}$  can, in theory, be realised with a new single form  $[U_{\pm\pm\pm}]$ , a single modifier  $U[\pm\pm\pm]$ , as a simple modification  $[U_{\pm\pm}]_{\pm}$  (on the basis of  $[U_{\pm\pm}]$ ), as modification on the basis of  $[U_{\pm}]$ , in which case modification may be reiterated  $[[U_{\pm}]_{\pm}]_{\pm}$ , or non-reiterated  $[U_{\pm}][\pm\pm]$ . Not all options are attested in our data. Specifically, we could find no structure that corresponds to  $[U_{\pm\pm\pm}]$  in our data. That is, there were no conventionalised forms that were based on lexicalisation higher than  $[U_{\pm\pm}]$ .

Table 8: Labelling of temporal reference points (by pattern)

	Based on U		Based on $U_{\pm}$		Based on $U_{\pm\pm}$ or higher	
<b><math>U_{\pm}</math></b>	Y[-]	<i>last year</i> (Eng.)	[D-]	<i>hier</i> (Fr.)	---	
<b><math>U_{\pm\pm}</math></b>	D[++]	<i>ngày kia</i> (Viet.) 'day a-bit-further'	[D+] +	<i>pasado mañana</i> (Sp.) 'passed D+1'	[D++]	<i>etzi</i> (Basque)

<b>U+++</b>	D[+++]	<i>ngày kia</i> (Viet.) 'day a-bit-further'	[[D+]+]+	<i>popojutrze</i> (Pol.) 'after.after.D+1'	[D--]-	<i>sakiototoi</i> (Jap.) 'before.D-2'
	---		[D+]++	<i>pasanfarda</i> (Farsi) 'thereafter.D+1'	[U+++]	---

Note that the modifiers used in [U-]- and [U+]+ constructions frequently enable attested productive recursion, although the frequency of use is reduced with the increase in number of reiterations. This is illustrated in (3), showing a x5 reiteration of the modifier [U-]-]

- (7) **GODOT:** *Avant-avant-avant-avant-avant-hier, je passe au loin devant un arbre et une route de campagne.*  
**'GODOT:** Before-before-before-before-before-yesterday, I passed by a tree and a country road in the distance.'  
 (Tribute to Beckett's *En attendant Godot*;  
[http://desencyclopedia.wikia.com/wiki/En\\_attendant\\_Godot](http://desencyclopedia.wikia.com/wiki/En_attendant_Godot)).

In some languages, this recursive pattern is preferred in the spoken language (e.g., Moldavian/Romanian), while in other languages, although frozen forms of the recursive pattern exist for the mid-range proximity ( $\pm 2-3$  and rarely for higher), these forms appear to be used less in present day, and speakers show a preference for the numerical compositional structures that are the only option from more distal reference (e.g., Estonian, Hebrew, Farsi (in the latter, the enumerated form was apparently borrowed from Arabic)).

## 7. Conclusion

Linguistic temporal reference is a complex task, and requires not only knowledge of the appropriate conventionalised expressions, but also reference to other linguistic and extra-linguistic information. And even then, the identification of the intended reference may not be absolutely clear. The present study highlights a number of factors operative in the linguistic use of calendar expressions, which must be taken into consideration in the theoretical study of temporal reference and in its formalisation, if it is to reflect actual linguistic use. Several aspects have been highlighted above:

- **Transparency** – Given a specific calendar, temporal identification formally requires three components: unit size, orientation, and distance from the central anchor. In language, constructions that refer to distal units ( $\geq \pm 3$  depending on language) are typically compositionally transparent in this way, but constructions indicating more proximal reference, particularly the 0,  $\pm 1$  range, may be more opaque, and these components may be entirely unrecoverable as separate elements.

- **Distance** - The temporal domain can be divided into 3 (overlapping or fuzzily bounded) zones around the deictic anchor of the present (S), which extend both to the future and the past. Reference to the central and proximal zones uses specialised constructions, while distal reference employs numerical (sequential) constructions, with no (linguistic) upper bound.
- **Referential (in)dependence** - Some temporal expressions (most notably reference with the fundamental units) can identify the intended temporal referent on their own, while others cannot function independently and require temporal cues from the rest of the clause (e.g., the parts of the day) or extra-linguistic information (like the identity of Day S for the use of days of the week). Interestingly, even this extra information cannot guarantee accurate identification of the intended time.
- **Relation to abstract calendars** – Linguistic use of time-keeping terms does not parallel abstract calendar systems. Different types of linguistic calendar expressions show different relations with the concepts and segmentations of the abstract calendar. This is particularly evident with proximal reference using named days of the week, which shows sensitivity to Day S, but not to the containing segment Week S. This highlights the fact that linguistic encoding of time is more complex, and one cannot assume simple direct reference to abstract calendar systems.

These aspects are interrelated. Thus, specialised constructions referring to proximal referents tend to be more opaque, while more distal reference employs productive numerical constructions, which are more compositionally transparent. The absence of temporal information (such as orientation or distance in the named days of the week) clearly affects the degree of dependence of the relevant expressions; temporal reference with these 'incomplete' expressions is more affected by the relations of the selected expressions to the relevant calendar cycles (as evident in the discussion of the named days of the week above).

The data indicate that a calendar representation based on a one-to-one correspondence with the linear order of  $\mathbb{Z}$ , the set of integers, does not capture the linguistic use of deictic calendar segmentation. We therefore argue that calendar segmentations are not represented on a single super-sequence. Rather, we propose a basic tripartite structure of the localised calendar, which is composed of two ordinal sequences, the order of  $\mathbb{N}$ , the set of natural integers, linked through a central articulation. These sequences are (i) distinct from each other, to reflect non-equivalence of linguistic structures in the two orientations, and (ii) further divided into ranges, to reflect the impact of distance on the linguistic structure.

Our comparative study of linguistic calendar expressions is in its infancy. The results obtained so far reveal intriguing patterns. However, much more empirical work is necessary to establish the linguistic realisations of these patterns language-specifically and to determine their impact and prevalence cross-linguistically. There are additional aspects to be examined, such as increased specification of temporal reference, which may be interpreted as redundant, but could result from a sense of insufficient precision with some temporal linguistic expressions, as noted above for the named days of the week. Yet another aspect to be examined is speakers' choice among competing linguistic expressions for the same temporal reference, both in terms of granularity of unit, and among forms within the same granularity. The underlying hypothesis is that the choice of expression (especially in deixis) signals

not only temporal anchoring but also the speakers' subjective perception of the states of affairs at the time of utterance, including a subjective perception of distance.

## 8. References

- Abusch, D., 1997. Sequence of Tense and Temporal de re. *Linguistics and Philosophy* 20.1, 1-50.
- Ariel, M., 1990. *Accessing Noun-Phrase Antecedents*. Routledge, London.
- Arnauld, A., Lancelot C., 1660. *Grammaire générale et raisonnée*. Republication Paulet, paris.
- Aristotle, 1908. *Metaphysics*. Translated by William David Ross. [http://en.wikisource.org/wiki/Metaphysics/Book\\_V](http://en.wikisource.org/wiki/Metaphysics/Book_V).
- Bach E., 1981. On time, tense, and aspect: an essay in English metaphysics. In: Cole, P. (Ed.), *Radical Pragmatics*. Academic Press, New York, pp. 62–81.
- Battistelli, D., Minel, J.-L., Couto, J., Schwerm R.S., 2008. Representing and Visualizing Calendar Expressions in Texts. STEP '08 Proceedings of the 2008 Conference on Semantics in Text Processing. *Research in Computational Semantics* 1, 365–373.
- Beauzée, N., 1767. *Grammaire Générale*. Barbou, Paris.
- Benveniste, E., 1974. *Problèmes de Linguistique générale* 2. Gallimard, Paris, pp. 67-78.
- Berthonneau, A-M., 1989. *Composantes Linguistiques de la Référence Temporelle. Les Compléments de Temps, du Lexique a l'énoncé*. Doctoral thesis, University of Paris VII.
- Berthonneau, A-M., 1990. Site anaphorique et site diectique. Etude stratigraphique des compléments de temps. In: Kleiber, G., Tyvaert, J-E., (Eds.). *L'anaphore et ses domaines*. Klincksieck, Paris, pp. 1-50.
- Bhat, S., 1999. *The Prominence of Tense, Aspect and Mood*. John Benjamins, Amsterdam.
- Bonomo, A., 2001. The temporal systems of Dagaare and Dagbane: re-appraising the philosophy of linguistic relativity. *Journal of Cultural Studies* 3.1, 43-55.
- Borillo, A., 1983. Les adverbes de référence temporelle dans la phrase et dans le texte. *DRLAV* 29, 109-131.
- Borillo, A., 2005. Les adverbes temporels et la structuration du discours. In: Molendijk, A., Vet, C., (Eds.) *Temporalité et Attitude: Structuration du Discours et Expression de la Modalité*. *Cahiers Chronos* 12, 1-18.
- Botne R., Kershner, T., 2008. Tense and cognitive systems: on the organization of tense/aspect systems in Bantu languages and beyond. *Cognitive Linguistics* 19, 145-218.

- Bras, M., 2005. A propos de quelques noms de temps en occitan. In: Choi-Jonin, I., Bras, M., Rouquier, M., Dagnac, A., (Eds.). *Questions de Classification en Linguistique. Mélanges offerts au Professeur Christian Molinier*. Peter Lang, Berne, pp. 55-80.
- Brinton, L.J., Traugott, E.C., 2005. *Lexicalization and Language Change*. Cambridge University Press, Cambridge.
- Bull, W., 1960. *Time, Tense, and the Verb: A Study in Theoretical and Applied Linguistics, with Particular Attention to Spanish*. University of California Publications in Linguistics 19. University of California Press, Berkely.
- Cailliau, F., Giraudel, A., Arnulphy, B., 2009. Tracking Out-of-date newspaper Articles. *Advances in Computational Linguistics, Resaerch in Computing Science* 41, 277-288.
- Cauty, A., 2000. Numération a deux zéros chez les Mayas. *Repères* 41: 25-51.
- Chafe , W.L., 1976. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In: Li, C.N., (Ed.). *Subject and Topic*. Academic Press, New York, pp. 25-57.
- Chung, S., Timberlake, A., 1985. Tense, aspect and mood. In Shopen, T. (Ed.). *Language Typology and Syntactic Description*. Cambridge University Press, Cambridge, pp. 202-258.
- Comrie, B., 1985. *Tense*. Cambridge University Press, Cambridge.
- Cukierman, D., Delgrande, J., 1998. Expressing time intervals and repetition within a formalisation of calendars. *Computational Intelligence* 14, 563-597.
- Demirdache, H., Uribe-Etxebarria, M., 2007. The syntax of time arguments. *Lingua* 117.2, 330-366.
- Dahl, O., 1983. Temporal distance: remoteness distinctions in tense-aspect systems. *Linguistics* 21.1, 105-122.
- Dahl, O., 1985. *Tense and Aspect Systems*. Blackwell, Oxford.
- Dowty, D.R., 1982. Tenses, time adverbials and compositional semantic theory. *Linguistics and Philosophy* 5, 23-33.
- Evans, V., 2005. *The Structure of Time*. John Benjamins, Amsterdam.
- Filmore C., 1997. *Lectures on Deixis*. CSLI Publications, Stanford.
- Flickinger, D., 1996. English Time Expressions in an HPSG Grammar. In Gunji (edt), *Studies on the Iniversality of Constraint-based Phrase Strcuture Grammars*. Ministry Of Education, Science, and Culture, Japan. 1-8.
- Giorgi, A., 2008. Crosslinguistic variation and the syntax of tense. *University of Venice Working Papers in Linguistics* 18, 145-177.
- Giorgi, A., Pianesi, F., 1997. *Tense and Aspect: from Semantics to Morphosyntax*. Oxford University Press, New York.
- Giorgi, A., Pianesi, F., 2003. The day after: anaphoric temporal locutions. In: Guéron, J., Tasmowsky, L., (Eds.), *Perspective and Point of View*. University of Paris X Press, Paris, pp. 128-171.
- Girard, G., 1747. *Les Vrais Principes de la Langue Française*. Lebreton, Paris.

- Givon, T., 1972. Studies in Chibemba and Bantu Grammar. Studies in African Linguistics, Supplement 3.
- Gosselin, L., 1986. Sémantique de la Temporalité en Français. Un modèle Calculatoire et Cognitif du Temps et de L'aspect. Duculot, Louvain-la-Neuve.
- Gross, M., 2002. Les déterminants numéraux, un exemple: les dates horaires. *Langages* 145, 21-37.
- Guillaume, G., 1929. Temps et Verbe, Théorie des Aspects, des Modes et des Temps. Collection Linguistique publiée par la Société de Linguistique de Paris, 27.
- Gundel, J.K., Hedberg, N., Zacharski, R., 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69.2: 274–307.
- Haspelmath, M., 1997. From Space to Time. Temporal Adverbials in the World's Languages. Mouton de Gruyter, Berlin.
- Hornstein, N., 1993. As Time Goes By - Tense and Universal Grammar. MIT Press, Cambridge Ma.
- Jespersen, Otto. 1924. The Philosophy of Grammar. H. Holt and Company, New York.
- José, L., 2003. Noms temporels et adjectifs déictiques. Le problème de l'opposition *l'an dernier* vs *\*l'instant dernier*. *Cahiers Chronos* 11, 211-222.
- Kamp, H., Rohrer, C., 1983. Tense in texts. In: Bauerle, R., Schwarze, R., von Stechow, A. (Eds.). Meaning, Use and Interpretation of Language. de Gruyter, Berlin, pp. 250–269.
- Kleiber, G., 1993. Lorsque l'anaphore se lie aux temps grammaticaux. In: Vetters, C. (Ed.) *Le temps de la Phrase au Texte*, Presses Universitaires de Lille, Lille, 117-166
- Levinson, S.C., 2005. Deixis. In: Horn, L.R., Ward, G., (Eds.) *The Handbook of Pragmatics*, Blackwell, Oxford, pp. 97–122.
- Mazur, P., Dale, R., 2008. What's the Date? High Accuracy Interpretation of Weekday Names. In 22<sup>nd</sup> International Conference on Computational Linguistics Proceedings, 553-560.
- McTaggart, J.M.E., 1908. The unreality of time. *Mind* 68, 457–474.
- Nguyen, P., 1992. Vietnamese demonstratives revisited. *Mon-Khmer Studies* 20: 127-136.
- Pustejovsky, J., Ingria, R., Sauri, R., Castano, J., Littman J., Gaizauskas, R., Setzer, A., Katz, G., Mani, I., 2005a. The specification language Timeml. In: Mani, I., Pustejovsky, J., Gaizauskas, R., (Eds.). *The Language of Time: A Reader*. Oxford University Press, New York.
- Pustejovsky, J., Knippen, R., Littman, J., Sauri, R., 2005b. Temporal and event information in Natural Language Text. *Language Resources and Evaluation*. 39.2, 123-164.
- Reichenbach, H., 1947. Elements of Symbolic Logic. Free Press, New York.
- Saussure, L., de, 2003. Temps et Pertinence. Duculot, Bruxelles.

- Schwer, S., 2002. Formalizing Calendars with the Category of Ordinals. *Applied Intelligence* 17.3, 275-295.
- Smith, C., 1991. *The Parameter of Aspect*. Kluwer, Dordrecht.
- Smith, C., 2003. *Modes of Discourse*. Cambridge University Press, Cambridge.
- Smith, C., 2007. Tense and temporal interpretation. *Lingua* 117, 419-436.
- ter Meulen, A., 1995. *Representing Time in Natural Language. The Dynamic Interpretation of Tense and Aspect*. MIT Press, Cambridge, Ma.
- Traugott, E., 1978. On the expression of spatio-temporal relations. In: Greenberg, J.H., Moravcsik, A.H., (Eds.), *Universals of Human Language*, Stanford University Press, Stanford.
- Vendler, Z., 1967. Verbs and times. In: *Linguistics in Philosophy*. Cornell University Press, Ithaca, New-York, pp. 97-121.
- Vet, C., 1980. *Temps, Aspects et Adverbes de Temps en Français Contemporain : Essai de Sémantique Formelle*. Droz, Genève.
- Watson, J.C.E., 2002. *The Phonology and Morphology of Arabic*. Oxford University Press, Oxford.
- Wilmet, M., 1976. *Etudes de Morpho-syntaxe Verbal*. Klincksieck, Paris.
- Wilmet M., 2003. *Grammaire Critique du Français*. Duculot, Bruxelles.
- Zagona, K. 2007. Approaches to tense and tense construal. *Lingua* 117.2, 464-502.